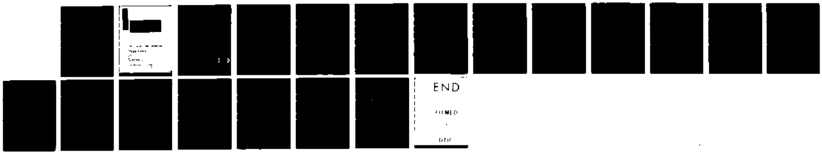
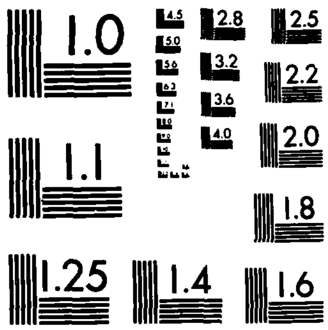
AD-A123 892 MINIMAX ESTIMATION OF A MULTIVARIATE NORMAL MEAN UNDER A QUADRATIC LOSS F..(U) FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS P LIN ET AL. JUN 82 1/1

UNCLASSIFIED FSU-STATISTICS-M633 N00014-80-C-0093 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MINIMAX ESTIMATION OF A MULTIVARIATE NORMAL MEAN
UNDER A QUADRATIC LOSS FUNCTION WITH UNKNOWN WEIGHTS

by

Pi-Erh Lin[1] and Amany M. Mousa

FSU Statistics Report No. M633
ONR Technical Report No. 158

June, 1982

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

[1]This work was supported by the Army, Navy and Air Force under Office of Naval Research Contract No. N00014-80-C-0093. Reproduction in whole or in part is permitted for any purpose of the United States Government.

MINIMAX ESTIMATION OF A MULTIVARIATE NORMAL MEAN UNDER A QUADRATIC LOSS FUNCTION WITH UNKNOWN WEIGHTS

by

Pi-Erh Lin[1] and Amany M. Mousa

Florida State University

SUMMARY

Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, $p \geq 3$, where $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_p^2)$ and let $s_i/\sigma_i^2 \sim \chi_{n_i}^2$ $(i = 1, \ldots, p)$, independent of $\underset{\sim}{X}$. This paper obtains a class of minimax estimators $\underset{\sim}{\delta}$ for $\underset{\sim}{\mu}$ subject to the loss function $(\underset{\sim}{\delta} - \underset{\sim}{\mu})'Q(\underset{\sim}{\delta} - \underset{\sim}{\mu})/\mathrm{tr}(Q\Sigma)$ where Q is a $p \times p$ diagonal matrix with unknown positive diagonal elements. It is assumed that an independent estimator $\hat{Q}^{-1}$ of Q^{-1} is available which satisfies certain conditions. The new minimax estimator is a function of $\underset{\sim}{X}$, s_i, and $\hat{Q}^{-1}$, and takes a form similar to the minimax estimator obtained by Berger and Bock (Ann. Statist. 4 (1976), 642-648) for the case when Q is known. A class of minimax estimators for $\underset{\sim}{\mu}$ is also obtained for the special case when $\sigma_1^2 = \ldots = \sigma_p^2$.

---

[1] This work was supported by the Army, Navy and Air Force under Office of Naval Research Contract No. N00014-80-C-0093. Reproduction in whole or in part is permitted for any purpose of the United States Government.

AMS 1970 subject classifications: Primary 62C07; Secondary 62F10, 62H99

Key words and phrases. Minimax, unknown variances, quadratic loss with unknown weights, risk function.

1. INTRODUCTION

Let $\underset{\sim}{X} = (X_1, \ldots, X_p)'$ be a p-variate normal random vector with mean $\underset{\sim}{\mu}$ and covariance matrix Σ. This paper considers the minimax estimation of $\underset{\sim}{\mu}$ by $\underset{\sim}{\delta}$ (delta) relative to a certain quadratic loss function with <u>unknown</u> weights. To the best of our knowledge, this is the first time in the literature a loss function of this type is considered in estimating $\underset{\sim}{\mu}$. The minimax estimation of $\underset{\sim}{\mu}$ relative to other types of quadratic loss functions has been extensively studied since Stein (1956) showed that the maximum likelihood estimator, $\underset{\sim}{X}$ (chi), is inadmissible, when $p \geq 3$, relative to the loss function given by a stated equation,

$$L_1(\underset{\sim}{\delta};\ \underset{\sim}{\mu}, \Sigma) = (\underset{\sim}{\delta} - \underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{\delta} - \underset{\sim}{\mu}). \tag{1.1}$$

In particular, minimax estimators for $\underset{\sim}{\mu}$ relative to (1.1) are obtained by Baranchik (1970) when Σ is known, by Strawderman (1973) when $\Sigma = \sigma^2 I_p$, and by Lin and Tsai (1973) and Efron and Morris (1976) when Σ is unknown, among other authors. In the 1970's, researchers shifted their concern to the minimax estimation of $\underset{\sim}{\mu}$ relative to the loss function given by

$$L_2(\underset{\sim}{\delta};\ \underset{\sim}{\mu}, \Sigma) = (\underset{\sim}{\delta} - \underset{\sim}{\mu})'Q(\underset{\sim}{\delta} - \underset{\sim}{\mu})/tr(Q\Sigma), \tag{1.2}$$

where $Q = diag(q_1, \ldots, q_p)$ under the assumption that $q_1, \ldots, q_p$ are <u>known positive constants</u>. Successful results are reported by Berger (1976a, b), Berger and Bock (1976), Berger <u>et al</u> (1977), and

Gleser (1979), among others. Recently, Lin and Mousa (1981a, b), in studying the minimax estimation under both loss functions, take a convex combination of (1.1) and (1.2) as the loss function and obtain classes of minimax estimators for $\underset{\sim}{\mu}$ under the assumption that Σ is an unknown positive definite diagonal matrix.

In this paper we will obtain a class of minimax estimators for $\underset{\sim}{\mu}$ relative to the loss function given by (1.2) where $q_1, \ldots, q_p$ are <u>unknown positive weights</u>. These positive weights should reflect the relative importance of the parameters to be estimated in such a way that the more important the component to be estimated is, the more weight should be placed on that component in the loss function. Usually, the parameters to be estimated are not equally important and, quite often, their relative importance is not completely known, resulting in an unknown assignment of weights. Since the real weight matrix Q is unknown, we will assume that an estimator $\hat{Q}$ of Q (or $\hat{Q}^{-1}$ of Q^{-1}) is available and that $\hat{Q}$ is independent of the underlying distribution. Examples in which Q is unknown are

(1) Let $\underset{\sim}{Z}_1, \ldots, \underset{\sim}{Z}_n$ be a random sample taken from $N_p(\underset{\sim}{\zeta}, \Gamma)$ where both $\underset{\sim}{\zeta}$ and Γ are unknown. Consider the problem of estimating $\underset{\sim}{\zeta}$ with respect to the loss $L_1(\underset{\sim}{\delta}; \underset{\sim}{\zeta}, \Gamma)$. Let $\bar{\underset{\sim}{Z}} = (1/n)\sum_{i=1}^{n} \underset{\sim}{Z}_i$. It is well known that $W = \sum_{i=1}^{n} (\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}})(\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}})'$ has a Wishart distribution with parameters n - 1 and Γ, independent of the $\underset{\sim}{Z}_i$'s. Thus, W/(n - 1) is an unbiased estimator of Γ.

(2) Let $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ be a random sample taken from $N_{p+1}(\underset{\sim}{\xi}, \Sigma)$. Partition $\underset{\sim}{X}_i$, $\underset{\sim}{\xi}$, and Σ as in Stein (1960), i.e.,

$$\underset{\sim}{X}_i = \begin{pmatrix} Y_i \\ \underset{\sim}{Z}_i \end{pmatrix}, \quad \underset{\sim}{\xi} = \begin{pmatrix} \eta \\ \underset{\sim}{\zeta} \end{pmatrix}, \text{ and } \quad \Sigma = \begin{pmatrix} A & \underset{\sim}{B}' \\ \underset{\sim}{B} & \Gamma \end{pmatrix},$$

where Y_i, η and A are 1×1, $\underset{\sim}{Z}_i$, $\underset{\sim}{\zeta}$ and $\underset{\sim}{B}$ are $p \times 1$ vectors, and Γ is an unknown $p \times p$ positive definite matrix. Then,

$$E(Y_i | \underset{\sim}{Z}_i) = \underset{\sim}{\beta}' \underset{\sim}{Z}_i + \alpha, \qquad i = 1, \ldots, p,$$

where

$$\underset{\sim}{\beta} = \Gamma^{-1} \underset{\sim}{B} \quad \text{and} \quad \alpha = \eta - \underset{\sim}{\beta}' \underset{\sim}{\zeta}.$$

Now, consider the problem of estimating the regression vector $\underset{\sim}{\beta}$ with respect to the loss

$$L(\hat{\beta}, \Sigma) = (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})' \Gamma (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) / (A - \underset{\sim}{B}' \Gamma^{-1} \underset{\sim}{B}). \qquad (1.3)$$

Note that $\Gamma / (A - \underset{\sim}{B}' \Gamma^{-1} \underset{\sim}{B})$ is unknown and it may be estimated by $V / (T - \underset{\sim}{U}' V^{-1} \underset{\sim}{U})$, where

$$\begin{pmatrix} T & \underset{\sim}{U}' \\ \underset{\sim}{U} & V \end{pmatrix} = \sum_{i=1}^{n} (\underset{\sim}{X}_i - \bar{\underset{\sim}{X}})(\underset{\sim}{X}_i - \bar{\underset{\sim}{X}})', \qquad \bar{\underset{\sim}{X}} = (1/n) \sum_{i=1}^{n} \underset{\sim}{X}_i.$$

(3) Consider the problem of estimating the means $\mu_1, \ldots, \mu_p$ $(p \geq 3)$ of certain products. Let $q_1, \ldots, q_p$ be the demands for the products which are unknown but which may be estimated, e.g., by some leading economic index or by some other method. The larger

the demand is for a product the more important it will be to estimate the mean product accurately so that its unit price may be fairly adjusted in the market. The use of loss function (1.2) with unknown weights in this context is not uncommon; in fact, it also can be found in psychology, education, and social research in which an independent source of information, that reflects the relative importance of the parameters to be estimated, is available.

The above examples suggest that, in practice, the matrix Q^{-1} need not be known rather that it can be estimated by $\hat{Q}^{-1}$ which has finite second moments.

In the following section, a new class of minimax estimators will be obtained for μ when the covariance matrix $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_p^2)$ with unknown σ_i^2, $i = 1, \ldots, p$. The special case when $\sigma_1^2 = \ldots = \sigma_p^2$ will be treated in Section 3.

2. MINIMAX ESTIMATION WHEN $\Sigma = \mathrm{DIAG}(\sigma_1^2, \ldots, \sigma_p^2)$

Let $\underset{\sim}{X} = (X_1, \ldots, X_p)'$ be a p-variate normal vector ($p \geq 3$) with unknown mean vector $\underset{\sim}{\mu} = (\mu_1, \ldots, \mu_p)'$ and unknown diagonal covariance matrix $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_p^2)$. Assume that $s_i/\sigma_i^2 \sim \chi^2_{n_i}$, $i = 1, \ldots, p$, and $s_1, \ldots, s_p$ are mutually independent, and are independent of $\underset{\sim}{X}$. It is desired to estimate the mean vector $\underset{\sim}{\mu}$ under the loss (1.2) where Q is a $p \times p$ diagonal matrix with unknown diagonal elements, $q_i > 0$, $i = 1, \ldots, p$. Let $\hat{Q}$ be a $p \times p$ diagonal random matrix with diagonal elements $\hat{q}_i$, $i = 1, \ldots, p$, independent of $\underset{\sim}{X}$ and s_i, such that for $i = 1, \ldots, p$,

$$E(q_i/\hat{q}_i) = c_i, \tag{2.1a}$$

and

$$E(q_i/\hat{q}_i)^2 = d_i, \tag{2.1b}$$

with c_i and d_i being known positive constants. Let

$$C = \mathrm{diag}(c_1, \ldots, c_p) \tag{2.2a}$$

and

$$D = \mathrm{diag}(d_1, \ldots, d_p). \tag{2.2b}$$

Assume that

$$\beta = ch_{max}(DQ^{-1}) \tag{2.3}$$

is known. Define $w_i = s_i/(n_i - 2)$, $i = 1, \ldots, p$, and let $W = \text{diag}(w_1, \ldots, w_p)$. Then W is independent of $\underset{\sim}{X}$ and $\hat{Q}$.

For an estimator $\underset{\sim}{\delta}(\underset{\sim}{X}, W, \hat{Q})$ let $R(\underset{\sim}{\delta}; \underset{\sim}{\mu}, \Sigma) = E_{W,\underset{\sim}{X},\hat{Q}} L_2(\underset{\sim}{\delta}; \underset{\sim}{\mu}, \Sigma)$ denote the risk function. It is easily verified that the maximum likelihood estimator, $\underset{\sim}{X}$, is minimax relative to the loss function given by (1.2) with constant risk equal to 1. Thus an estimator $\underset{\sim}{\delta}(\underset{\sim}{X}, W, \hat{Q})$ will be minimax if and only if $R(\underset{\sim}{\delta}; \underset{\sim}{\mu}, \Sigma) \le 1$ for all $\underset{\sim}{\mu}$ and Σ. In this section, the estimator $\underset{\sim}{\delta}(\underset{\sim}{X}, W, \hat{Q})$ given by (2.4) will be shown to be minimax by proving that $R(\underset{\sim}{X}; \underset{\sim}{\mu}, \Sigma) - R(\underset{\sim}{\delta}; \underset{\sim}{\mu}, \Sigma) \ge 0$ for all $\underset{\sim}{\mu}$ and Σ. The following lemmas are useful in evaluating the difference in risks; they are stated here without proof.

Lemma 2.1. [Stein (1974, 1981)]. Let $Y \sim N(0, 1)$ and let g be an absolutely continuous function, $g\colon R \to R$ such that $g(y)\exp(-y^2/2) \to 0$ as $y \to \pm\infty$. Then

$$E_Y[g'(Y)] = E_Y[Yg(Y)].$$

Lemma 2.2. [Efron and Morris (1976)]. Let $U \sim \chi_n^2$ and let g be an absolutely continuous function, $g\colon R^+ \to R^+$ such that $g(u)u^{n/2}\exp(-u/2) \to 0$ as $u \to 0^+$ or as $u \to \infty$. Then

$$E_U[Ug(U)] = nE_U[g(U)] + 2E_U[Ug'(U)].$$

Corollary 2.2.1. [Berger and Bock (1976)]. Let U and g be as defined in Lemma 2.2. Let $Z = cU/(n - 2)$, $c > 0$, and $h(Z) = g[(n - 2)Z/c]$. Then

$$E_Z[(n - 2)Zh(Z)/c] = nE_Z[h(Z)] + 2E_Z[Zh'(Z)].$$

<u>Lemma 2.3.</u> [Lehmann (1966)]. Let S be any random variable, and let $p_1(S)$ and $p_2(S)$ map the real line into itself. If $p_1(S)$ and $p_2(S)$ are either both nonincreasing in S or both nondecreasing in S, then

$$E_S[p_1(S)p_2(S)] \geq E_S[p_1(S)]E_S[p_2(S)].$$

The above lemmas have been frequently used to establish the minimaxity of an estimator for a multivariate normal mean; they are included here for ease of reference.

The following theorem will prove the minimaxity of $\underset{\sim}{\delta}(\underset{\sim}{X}, W, \hat{Q})$ given by (2.4). In the theorem we will use tr(A) and $ch_{max}(A)$ (or $ch_{min}(A)$) to denote the trace and the maximum (or minimum) characteristic root of a square matrix A. As in Berger and Bock (1976), let $T = \min_{1\leq i\leq p}(\chi^2_{n_i}/n_i)$ and $\tau = \tau(n_1, \ldots, n_p) = E(T^{-1})$.

<u>Theorem 2.4.</u> The estimator

$$\underset{\sim}{\delta}(\underset{\sim}{X}, W, \hat{Q}) = [I_p - r(\underset{\sim}{X}, W)\|\underset{\sim}{X}\|_W^{-2}\hat{Q}^{-1}W^{-1}]\underset{\sim}{X}, \tag{2.4}$$

where $\|\underset{\sim}{X}\|_W^2 = \sum_{i=1}^{p} X_i^2/w_i^2$, is minimax for $\underset{\sim}{\mu}$, provided that

(i) $0 \leq r(\underset{\sim}{X}, W) \leq 2[tr(C) - 2\tau\ ch_{max}(C)]/\beta$

with $tr(C) \geq 2\tau\ ch_{max}(C)$,

(ii) $r(\underset{\sim}{X}, W)$ is nondecreasing in $|X_i|$, $i = 1, \ldots, p$,

(iii) $r(\underset{\sim}{X}, W)$ is nonincreasing in w_i, $i = 1, \ldots, p$, and

(iv) $r(\underset{\sim}{X}, W)\|\underset{\sim}{X}\|_W^{-2}$ is nondecreasing in w_i, $i = 1, \ldots, p$.

<u>Proof</u>. Write $r = r(\underset{\sim}{X}, W, \hat{Q})$. Then the difference between the risk of $\underset{\sim}{X}$ and that of $\underset{\sim}{\delta}$, except for the factor $2/\mathrm{tr}(Q\Sigma)$, is

$$E_{W,\underset{\sim}{X},\hat{Q}}\left[\frac{r(\underset{\sim}{X}-\underset{\sim}{\mu})'Q\hat{Q}^{-1}W^{-1}\underset{\sim}{X}}{\|\underset{\sim}{X}\|_W^2} - \frac{r^2\underset{\sim}{X}'W^{-1}\hat{Q}^{-1}Q\hat{Q}^{-1}W^{-1}\underset{\sim}{X}}{2\|\underset{\sim}{X}\|_W^4}\right]$$

$$= E_{W,\underset{\sim}{X},\hat{Q}}\left[\frac{r}{\|\underset{\sim}{X}\|_W^2}\sum_{i=1}^{p}\frac{X_i(X_i-\mu_i)q_i}{\hat{q}_i w_i} - \frac{r^2}{2\|\underset{\sim}{X}\|_W^4}\sum_{i=1}^{p}\frac{X_i^2 q_i}{w_i^2\hat{q}_i^2}\right]$$

$$= E_{W,\underset{\sim}{X}}\left[\frac{r}{\|\underset{\sim}{X}\|_W^2}\sum_{i=1}^{p}\frac{X_i(X_i-\mu_i)c_i}{w_i} - \frac{r^2}{2\|\underset{\sim}{X}\|_W^4}\sum_{i=1}^{p}\frac{X_i^2 d_i}{w_i^2 q_i}\right]. \qquad (2.5)$$

The minimaxity of $\underset{\sim}{\delta}(\underset{\sim}{X}, W, \hat{Q})$ will be established if we can show that (2.5) ≥ 0. To evaluate the expectation in (2.5) we proceed as follows. The first term in the last expression of (2.5) becomes

$$E_{W,\underset{\sim}{X}}\left[\frac{r}{\|\underset{\sim}{X}\|_W^2}\sum_{i=1}^{p}\frac{\sigma_i c_i X_i}{w_i}\left(\frac{X_i-\mu_i}{\sigma_i}\right)\right]$$

$$= E_W\left\{E_{\underset{\sim}{X}|W}\left[\sum_{i=1}^{p}\frac{\sigma_i^2 c_i}{\|\underset{\sim}{X}\|_W^2 w_i}\left(r - \frac{2rX_i^2}{\|\underset{\sim}{X}\|_W^2 w_i^2} + X_i\frac{\partial r}{\partial X_i}\right)\right]\right\}$$

$$\geq E_{W,\underset{\sim}{X}}\left[\sum_{i=1}^{p}\frac{\sigma_i^2 c_i}{\|\underset{\sim}{X}\|_W^2 w_i}\left(r - \frac{2rX_i^2}{\|\underset{\sim}{X}\|_W^2 w_i^2}\right)\right]. \qquad (2.6)$$

The equality in (2.6) is obtained by an application of Lemma 2.1 with $y_i = (X_i - \mu_i)/\sigma_i$ and $g(y_i) = r\|\underset{\sim}{X}\|_W^{-2}\sigma_i c_i(\sigma_i y_i + \mu_i)/w_i$, and the inequality follows from Assumption (ii). Now let $U_i = (n_i - 2)w_i/\sigma_i^2$ and $h(w_i) = r/(\|\underset{\sim}{X}\|_W^2 w_i)$, $i = 1, \ldots, p$. Then by an application of Corollary 2.2.1, it follows that

$$E_{W,\underset{\sim}{X}}\left[\frac{r\sigma_i^2 c_i}{\|\underset{\sim}{X}\|_W^2 w_i}\right]$$

$$= E_{W,\underset{\sim}{X}}\left[\frac{rc_i}{\|\underset{\sim}{X}\|_W^2} - \frac{4r\sigma_i^2 c_i X_i^2}{(n_i - 2)\|\underset{\sim}{X}\|_W^4 w_i^3}\right.$$

$$\left. - \frac{2\sigma_i^2 c_i}{(n_i - 2)\|\underset{\sim}{X}\|_W^2}\left(\frac{\partial r}{\partial w_i}\right)\right] . \qquad (2.7)$$

Note that identity (2.7) is similar to (2.3) of Berger and Bock (1976) with their $q_1 = \ldots = q_p = 1$.

Substituting (2.6) and (2.7) into (2.5) and noting that $(\partial r/\partial w_i) \le 0$ for $i = 1, \ldots, p$, by Assumption (iii), the last expression of (2.5) is bounded by

$$E_{W,\underset{\sim}{X}}\left[\frac{r\,tr(C)}{\|\underset{\sim}{X}\|_W^2} - \frac{2r}{\|\underset{\sim}{X}\|_W^2}\sum_{i=1}^{p}\frac{n_i\sigma_i^2 X_i^2 c_i}{(n_i-2)w_i^3} - \frac{r^2}{2\|\underset{\sim}{X}\|_W^4}\sum_{i=1}^{p}\frac{X_i^2 d_i}{w_i^2 q_i}\right]$$

$$\geq E_{W,\underset{\sim}{X}}\left[\frac{r\,tr(C)}{\|\underset{\sim}{X}\|_W^2} - \frac{2r\,ch_{max}(C)}{\|\underset{\sim}{X}\|_W^4}\max_{1\le i\le p}\left(\frac{n_i\sigma_i^2}{(n_i-2)w_i}\right)\sum_{i=1}^{p}\frac{X_i^2}{w_i^2} - \frac{r^2}{2\|\underset{\sim}{X}\|_W^4}\sum_{i=1}^{p}\frac{X_i^2 d_i}{w_i^2 q_i}\right]$$

$$\geq E_{W,\underset{\sim}{X}}\left[\frac{r\,tr(C)}{\|\underset{\sim}{X}\|_W^2} - \frac{2r g(W)\,ch_{max}(C)}{\|\underset{\sim}{X}\|_W^2} - \frac{r^2\beta}{2\|\underset{\sim}{X}\|_W^2}\right]$$

$$= E_{W,\underset{\sim}{X}}\left\{(r\|\underset{\sim}{X}\|_W^{-2})\left[tr(C) - 2g(W)\,ch_{max}(C) - \frac{r\beta}{2}\right]\right\}, \tag{2.8}$$

where $g(W) = \max_{1\le i\le p}\left(\frac{n_i\sigma_i^2}{(n_i-2)w_i}\right)$ is nonincreasing in w_i, $i = 1, \ldots, p$, and $\beta = ch_{max}(DQ^{-1})$.

Now (2.5) ≥ 0 if we can show that (2.8) ≥ 0. But Lemma 2.3 implies that the last expression of (2.8) is further bounded below by

$$E_{\underset{\sim}{X}}\left\{E_W(r\|\underset{\sim}{X}\|_W^{-2})\left[tr(C) - 2E_W g(W)\,ch_{max}(C) - \frac{\beta}{2}E_W(r)\right]\right\}$$

$$= E_{X}\left\{E_W(r\|\underset{\sim}{X}\|_W^{-2})\left[tr(C) - 2\tau\,ch_{max}(C) - \frac{\beta}{2}E_W(r)\right]\right\}$$

which is nonnegative by Assumption (i) of Theorem 2.4. This establishes that (2.5) ≥ 0 and the theorem is proved. □

Note that the theorem will be vacuous unless $2\tau \le tr(C)/ch_{max}(C)$. The values of τ can be calculated from a formula, due to Berger and Bock (1976), if the n_i's are even. Recall that $\tau = E\left[\max_{1\le i\le p}(n_i/\chi^2_{n_i})\right]$

Then,

$$\tau = \sum_{k=1}^{p} \sum \left\{ \left[\prod_{i \neq k} \frac{m_i^{(m_i - j(i))}}{(m_i - j(i))!} \right] \frac{(m - J(k) - 2)!\, m_k^{m_k}}{(m_k - 1)!\, m^{(m - J(k) - 1}} \right\},$$

where $m_i = n_i/2$, $m = \sum_{i=1}^{p} m_i$, $J(k) = \sum_{i \neq k} j(i)$, and the inner summation is over all combinations $[j(1), j(2), \ldots, j(k-1), j(k+1), \ldots, j(p)]$ where the $j(i)$ are integers between 1 and m_i, inclusive. On the other hand, for the important special case when $C = cI_p$, $c > 0$, $tr(C)/[2ch_{max}(C)] = p/2$. In this case, $2\tau \leq p$, which has been shown by Berger and Bock (1976) to hold when $p \geq 3$ and $n_i \geq N$ for some large positive integer N.

Remark. The assumption that $\beta = ch_{max}(DQ^{-1})$ is known seems undesirable. However, since the loss (1.2) is invariant with respect to a same positive scale change on all elements of Q, we may without loss of generality assume that $ch_{min}(Q) = 1$ in (1.2). Thus $\beta \leq ch_{max}(D)$ and Assuption (i) of Theorem 2.4 may be replaced by

(i′) $0 \leq r(\underset{\sim}{X}, W) \leq 2[tr(C) - 2\tau ch_{max}(C)]/ch_{max}(D)$, with $tr(C) \geq 2\tau ch_{max}(C)$.

3. MINIMAX ESTIMATION WHEN $\Sigma = \sigma^2 I_p$.

In this section we will consider the special case when $\sigma_1^2 = \dots = \sigma_p^2 = \sigma^2$, say. Specifically, let $\underset{\sim}{X} = (X_1, \dots, X_p)'$ be a p-variate normal vector ($p \geq 3$) with unknown mean vector $\underset{\sim}{\mu}$ and unknown covariance matrix $\Sigma = \sigma^2 I_p$. Assume that $s/\sigma^2 \sim \chi_n^2$ ($n \geq 3$), independent of $\underset{\sim}{X}$. It is desired to estimate the mean vector $\underset{\sim}{\mu}$ under the loss (1.2) where Q is a $p \times p$ diagonal matrix with unknown diagonal elements $q_i > 0$, $i = 1, \dots, p$. Let $\hat{Q}$ be as defined in Section 2 satisfying conditions (2.1) and (2.2). Define $w = s/(n-2)$. Then w is independent of $\underset{\sim}{X}$ and $\hat{Q}$. In the following theorem, an estimator $\underset{\sim}{\delta}$ which is a function of $\underset{\sim}{X}$, w, and $\hat{Q}$ is shown to be minimax with $R(\underset{\sim}{\delta}; \underset{\sim}{\mu}, \sigma^2) \leq 1$ for all $\underset{\sim}{\mu}$ and σ^2.

Theorem 3.1. The estimator

$$\underset{\sim}{\delta}(\underset{\sim}{X}, w, \hat{Q}) = [I_p - r(\underset{\sim}{X}, w)\|\underset{\sim}{X}\|_w^{-2}\hat{Q}^{-1}w^{-1}]\underset{\sim}{X},$$

where $\|\underset{\sim}{X}\|_w^2 = \underset{\sim}{X}'\underset{\sim}{X}/w^2$, is minimax for the mean vector $\underset{\sim}{\mu}$, provided that

(i) $0 \leq r(\underset{\sim}{X}, w) \leq 2\left(\frac{n-2}{n+2}\right)[tr(C) - 2\, ch_{max}(C)]/\beta$, with $tr(C) \geq 2\, ch_{max}(C)$, $n \geq 3$, and $\beta = ch_{max}(DQ^{-1})$,

(ii) $r(\underset{\sim}{X}, w)$ is nondecreasing in $|X_i|$, $i = 1, \dots, p$,

(iii) $r(\underset{\sim}{X}, w)$ is nonincreasing in w, and

(iv) $r(\underset{\sim}{X}, w)\|\underset{\sim}{X}\|_w^{-2}$ is nondecreasing in w.

Proof. Write $r = r(\underset{\sim}{X}, w)$. Then the difference between the risk of $\underset{\sim}{X}$ and that of $\underset{\sim}{\delta}$, except for the factor $2/\mathrm{tr}(Q\Sigma)$, is

$$E_{w,\underset{\sim}{X},\hat{Q}}\left[\frac{r(\underset{\sim}{X}-\underset{\sim}{\mu})'Q\hat{Q}^{-1}\underset{\sim}{X}}{\|\underset{\sim}{X}\|_w^2 w} - \frac{r^2\underset{\sim}{X}'\hat{Q}^{-1}Q\hat{Q}^{-1}\underset{\sim}{X}}{2\|\underset{\sim}{X}\|_w^4 w^2}\right]$$

$$= E_{w,\underset{\sim}{X},\hat{Q}}\left[\frac{r}{\|\underset{\sim}{X}\|_w^2 w}\sum_{i=1}^{p}\frac{X_i(X_i-\mu_i)q_i}{\hat{q}_i} - \frac{r^2}{2\|\underset{\sim}{X}\|_w^4 w^2}\sum_{i=1}^{p}\frac{X_i^2 q_i}{\hat{q}_i^2}\right]$$

$$= E_{w,\underset{\sim}{X}}\left[\frac{r}{\|\underset{\sim}{X}\|_w^2 w}\sum_{i=1}^{p}X_i(X_i-\mu_i)c_i - \frac{r^2}{2\|\underset{\sim}{X}\|_w^4 w^2}\sum_{i=1}^{p}\frac{X_i^2 d_i}{q_i}\right], \tag{3.1}$$

which is similar to (2.5). To prove that $\underset{\sim}{\delta}$ is minimax, it is sufficient to show that (3.1) is ≥ 0, for all $\underset{\sim}{\mu}$ and σ^2. This may be established by simply modifying the proof of (2.5). The details are omitted. □

The remark given at the end of Section 2 applies to this case as well. Thus Assumption (i) of Theorem 3.1 may be replaced by

(i′) $0 \leq r(\underset{\sim}{X}, w) \leq 2\left(\frac{n-2}{n+2}\right)[\mathrm{tr}(C) - 2\,\mathrm{ch}_{\max}(C)]/\mathrm{ch}_{\max}(D)$,

where $\mathrm{tr}(C) \geq 2\,\mathrm{ch}_{\max}(C)$ and $n \geq 3$.

The theorem is not vacuous since for the special case when $C = cI_p$, $c > 0$, the upper bound of $r(\underset{\sim}{X}, w)$ reduces to $2[(n-2)/(n+2)]c(p-2)/\mathrm{ch}_{\max}(D)$ which is nonnegative for all $n \geq 3$ and $p \geq 3$.

REFERENCES

BARANCHIK, A.J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. Ann. Math. Statist. 41, 642-645.

BERGER, J. (1976a). Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss. Ann. Statist. 4, 223-226.

BERGER, J. (1976b). Minimax estimation of a multivariate normal mean with arbitrary quadratic loss. Ann. Statist. 4, 256-264.

BERGER, J. and BOCK, M.E. (1976). Combining independent normal mean estimation problems with unknown variances. Ann. Statist. 4, 642-648.

BERGER, J., BOCK, M.E., BROWN, L.D., CASELLA, G., and GLESER, L. (1977). Minimax estimation of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix. Ann. Statist. 5, 763-771.

EFRON, B. and MORRIS, C. (1976). Families of minimax estimators of the mean of a multivariate normal distribution. Ann. Statist. 4, 11-21.

GLESER, L. (1979). Minimax estimation of a normal mean vector when the covariance matrix is unknown. Ann. Statist. 7, 838-846.

HUDSON, M. (1978). A natural identity for exponential families with applications in multiparameter estimation. Ann. Statist. 6, 473-484.

LEHMANN, D.L. (1966). Some concepts of dependence. Ann. Math. Statist. 37, 1137-1153.

LIN, P.E. and MOUSA, A. (1981a). Minimax estimation of a multivariate normal mean under a convex loss function. FSU Statistics Report No. 591.

LIN, P.E. and MOUSA, A. (1981b). Proper Bayes minimax estimators for a multivariate normal mean with unknown common variance under a convex loss function. FSU Statistics Report No. 593.

LIN, P.E. and TSAI, H.L. (1973). Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix. Ann. Statist. 1, 142-145.

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Proc. Third Berkeley Symp. Math. Statist. Prob. 1, 197-206, Univ. of California Press.

STEIN, C. (1960). Multiple regression. Contributions to Probability and Statistics (I. Olkin ed.). Stanford Univ. Press.

STEIN, C. (1974). Estimation of the parameters of a multivariate normal distribution, Part I. Tech. Report No. 63, Stanford University.

STEIN, C. (1981). Estimation of the mean of a multivariate normal distribution. Ann. Statist. 9, 1135-1151.

STRAWDERMAN, W. (1973). Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances. Ann. Statist. 1, 1189-1194.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER FSU Statistics M633 ONR Technical 158 — 2. GOVT ACCESSION NO. AD-A123 892	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and subtitle) Minimax Estimation of a Multivariate Normal Mean Under a Quadratic Loss Function With Unknown Weights	5. TYPE OF REPORT & PERIOD COVERED Technical 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Pi-Erh Lin and Amany M. Mousa	8. CONTRACT OR GRANT NUMBER(s) ONR Contract N00014-80-C-0093
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Florida State University Tallahassee, FL 32306	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Arlington, VA 22217	12. REPORT DATE June, 1982 13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this report)

Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS

Minimax, unknown variances, quadratic loss with unknown weights, risk function

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, $p \geq 3$, where $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_p^2)$ and let $s_i/\sigma_i^2 \sim \chi^2_{n_i}$ $(i = 1, \ldots, p)$, independent of $\underset{\sim}{X}$. This paper obtains a class of minimax estimators $\underset{\sim}{\delta}$ for $\underset{\sim}{\mu}$ subject to the loss function $(\underset{\sim}{\delta} - \underset{\sim}{\mu})'Q(\underset{\sim}{\delta} - \underset{\sim}{\mu})/\mathrm{tr}(Q\Sigma)$ where Q is a $p \times p$ diagonal matrix with unknown positive diagonal elements. It is assumed that an independent estimator $\hat{Q}^{-1}$ of Q^{-1} is available which satisfies certain conditions. The new minimax estimator is a function of $\underset{\sim}{X}$, s_i, and $\hat{Q}^{-1}$, and takes a form similar to the minimax estimator obtained by Berger and Bock (Ann. Statist. 4 (1976), 642-648) for the case when Q is known. A class of minimax estimators for $\underset{\sim}{\mu}$ is also obtained for the special case when $\sigma_1^2 = \ldots = \sigma_p^2$.